

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/560,250  
Source: TFW0  
Date Processed by STIC: 07/31/2006

# ***ENTERED***



IFWO

## RAW SEQUENCE LISTING

DATE: 07/31/2006

PATENT APPLICATION: US/10/560,250

TIME: 08:48:16

Input Set : N:\RJAVED\10560250.txt

Output Set: N:\CRF4\07312006\J560250.raw

4 <110> APPLICANT: The Scripps Research Institute  
 5 The Regents of the University of California  
 6 Wu, Eugene  
 7 Nemerow, Glen R.  
 8 Stewart, Phoebe  
 10 <120> TITLE OF INVENTION: MODIFIED FIBER PROTEINS FOR EFFICIENT  
 11 RECEPTOR BINDING  
 13 <130> FILE REFERENCE: 22908-1237PC  
 C--> 15 <140> CURRENT APPLICATION NUMBER: US/10/560,250  
 C--> 16 <141> CURRENT FILING DATE: 2005-12-12  
 18 <150> PRIOR APPLICATION NUMBER: not assigned  
 19 <151> PRIOR FILING DATE: <151 2003-06-11  
 21 <160> NUMBER OF SEQ ID NOS: 70  
 23 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
 25 <210> SEQ ID NO: 1  
 26 <211> LENGTH: 48  
 27 <212> TYPE: DNA  
 28 <213> ORGANISM: Artificial Sequence  
 30 <220> FEATURE:  
 31 <223> OTHER INFORMATION: primer  
 33 <400> SEQUENCE: 1  
 34 tgtcttgaat ccaagatgaa gcgcgcccgc cccagcgaag atgacttc 48  
 36 <210> SEQ ID NO: 2  
 37 <211> LENGTH: 48  
 38 <212> TYPE: DNA  
 39 <213> ORGANISM: Artificial Sequence  
 41 <220> FEATURE:  
 42 <223> OTHER INFORMATION: primer  
 44 <400> SEQUENCE: 2  
 45 tggagctgggt gtggtccaca aagtgcgcgt gtcatttct gggttcca 48  
 47 <210> SEQ ID NO: 3  
 48 <211> LENGTH: 24  
 49 <212> TYPE: DNA  
 50 <213> ORGANISM: Artificial Sequence  
 52 <220> FEATURE:  
 53 <223> OTHER INFORMATION: primer  
 55 <400> SEQUENCE: 3  
 56 actttgtgga ccacaccagc tcca 24  
 58 <210> SEQ ID NO: 4  
 59 <211> LENGTH: 30  
 60 <212> TYPE: DNA  
 61 <213> ORGANISM: Artificial Sequence  
 63 <220> FEATURE:

## RAW SEQUENCE LISTING

DATE: 07/31/2006

PATENT APPLICATION: US/10/560,250

TIME: 08:48:16

Input Set : N:\RJAVED\10560250.txt

Output Set: N:\CRF4\07312006\J560250.raw

```

64 <223> OTHER INFORMATION: primer
66 <400> SEQUENCE: 4
67 cataacgcgg ccgcttcttt attcttgggc 30
69 <210> SEQ ID NO: 5
70 <211> LENGTH: 42
71 <212> TYPE: DNA
72 <213> ORGANISM: Artificial Sequence
74 <220> FEATURE:
75 <223> OTHER INFORMATION: primer
77 <400> SEQUENCE: 5
78 gtgctactaa acaattcctt cctggatcca gaattattgga ac 42
80 <210> SEQ ID NO: 6
81 <211> LENGTH: 42
82 <212> TYPE: DNA
83 <213> ORGANISM: Artificial Sequence
85 <220> FEATURE:
86 <223> OTHER INFORMATION: primer
88 <400> SEQUENCE: 6
89 gttccaatat tctggatcca ggaaggaatt gtttagtagc ac 42
91 <210> SEQ ID NO: 7
92 <211> LENGTH: 30
93 <212> TYPE: DNA
94 <213> ORGANISM: Artificial Sequence
96 <220> FEATURE:
97 <223> OTHER INFORMATION: primer
99 <400> SEQUENCE: 7
100 atgggatcca agatgaagcg cgcaagaccg 30
102 <210> SEQ ID NO: 8
103 <211> LENGTH: 30
104 <212> TYPE: DNA
105 <213> ORGANISM: Artificial Sequence
107 <220> FEATURE:
108 <223> OTHER INFORMATION: primer
110 <400> SEQUENCE: 8
111 tgggtgtggtc cacaaagtta gcttatcatt 30
113 <210> SEQ ID NO: 9
114 <211> LENGTH: 48
115 <212> TYPE: DNA
116 <213> ORGANISM: Artificial Sequence
118 <220> FEATURE:
119 <223> OTHER INFORMATION: primer
121 <400> SEQUENCE: 9
122 aagctaactt tgtggaccac accagacaca tctccaaact gcacaatt 48
124 <210> SEQ ID NO: 10
125 <211> LENGTH: 28
126 <212> TYPE: DNA
127 <213> ORGANISM: Artificial Sequence
129 <220> FEATURE:
130 <223> OTHER INFORMATION: primer

```

## RAW SEQUENCE LISTING

DATE: 07/31/2006

PATENT APPLICATION: US/10/560,250

TIME: 08:48:16

Input Set : N:\RJAVED\10560250.txt

Output Set: N:\CRF4\07312006\J560250.raw

```

132 <400> SEQUENCE: 10
133 aaacacggcg gccgctcttt cattcttg                28
135 <210> SEQ ID NO: 11
136 <211> LENGTH: 45
137 <212> TYPE: DNA
138 <213> ORGANISM: Artificial Sequence
140 <220> FEATURE:
141 <223> OTHER INFORMATION: primer
143 <400> SEQUENCE: 11
144 ctttgtggac cacaccagac actagtccaa actgcacaat tgctc        45
146 <210> SEQ ID NO: 12
147 <211> LENGTH: 45
148 <212> TYPE: DNA
149 <213> ORGANISM: Artificial Sequence
151 <220> FEATURE:
152 <223> OTHER INFORMATION: primer
154 <400> SEQUENCE: 12
155 gagcaattgt gcagtttggga ctagtgtctg gtgtgggtcca caaag      45
157 <210> SEQ ID NO: 13
158 <211> LENGTH: 48
159 <212> TYPE: DNA
160 <213> ORGANISM: Artificial Sequence
162 <220> FEATURE:
163 <223> OTHER INFORMATION: primer
165 <400> SEQUENCE: 13
166 gcttaggtta acctcaagct ttttcttggt ttttttgaga ggtgggct    48
168 <210> SEQ ID NO: 14
169 <211> LENGTH: 48
170 <212> TYPE: DNA
171 <213> ORGANISM: Artificial Sequence
173 <220> FEATURE:
174 <223> OTHER INFORMATION: primer
176 <400> SEQUENCE: 14
177 agcccacctc tcaaaaaaac caggaaaaag cttgaggtta acctaagc    48
179 <210> SEQ ID NO: 15
180 <211> LENGTH: 72
181 <212> TYPE: DNA
182 <213> ORGANISM: Artificial Sequence
184 <220> FEATURE:
185 <223> OTHER INFORMATION: primer
187 <400> SEQUENCE: 15
188 atcagtatta acttgcagtg gagccttagg gtttacagtt aggcttccgg cctcgtccag 60
189 agagaggccg tt                                           72
191 <210> SEQ ID NO: 16
192 <211> LENGTH: 72
193 <212> TYPE: DNA
194 <213> ORGANISM: Artificial Sequence
196 <220> FEATURE:
197 <223> OTHER INFORMATION: primer

```

## RAW SEQUENCE LISTING

DATE: 07/31/2006

PATENT APPLICATION: US/10/560,250

TIME: 08:48:16

Input Set : N:\RJAVED\10560250.txt

Output Set: N:\CRF4\07312006\J560250.raw

```

199 <400> SEQUENCE: 16
200 ggaagcctaa ctgtaaaccc taaggctcca ctgcaagtta atactgattc aaacataaac 60
201 ctggaaatat ct 72
203 <210> SEQ ID NO: 17
204 <211> LENGTH: 72
205 <212> TYPE: DNA
206 <213> ORGANISM: Artificial Sequence
208 <220> FEATURE:
209 <223> OTHER INFORMATION: primer
211 <400> SEQUENCE: 17
212 atcattgtca aatgtcaacc cttctcttgc tcttacattt ataccaatgt tgtaatcaaa 60
213 ttctaggcca tg 72
215 <210> SEQ ID NO: 18
216 <211> LENGTH: 72
217 <212> TYPE: DNA
218 <213> ORGANISM: Artificial Sequence
220 <220> FEATURE:
221 <223> OTHER INFORMATION: primer
223 <400> SEQUENCE: 18
224 attggtataa atgtaagagc aagagaaggg ttgacatttg acaatgatgg tgccattaca 60
225 gtaggaaaca aa 72
227 <210> SEQ ID NO: 19
228 <211> LENGTH: 38
229 <212> TYPE: DNA
230 <213> ORGANISM: Artificial Sequence
232 <220> FEATURE:
233 <223> OTHER INFORMATION: primer
235 <400> SEQUENCE: 19
236 ctggacgagg ccggcagcct aactgtaaac cctaaggg 38
238 <210> SEQ ID NO: 20
239 <211> LENGTH: 38
240 <212> TYPE: DNA
241 <213> ORGANISM: Artificial Sequence
243 <220> FEATURE:
244 <223> OTHER INFORMATION: primer
246 <400> SEQUENCE: 20
247 gccttagggg ttacagttag gctgccggcc tcgtccag 38
249 <210> SEQ ID NO: 21
250 <211> LENGTH: 7960
251 <212> TYPE: DNA
252 <213> ORGANISM: Artificial Sequence
254 <220> FEATURE:
255 <223> OTHER INFORMATION: pDV67
257 <400> SEQUENCE: 21
258 gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60
259 ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
260 cgagcaaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
261 ttaggggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatagc cgttgacatt 240
262 gattattgac tagttattaa tagtaatcaa ttacgggggc attagttcat agcccatata 300

```

## RAW SEQUENCE LISTING

DATE: 07/31/2006

PATENT APPLICATION: US/10/560,250

TIME: 08:48:16

Input Set : N:\RJAVED\10560250.txt

Output Set: N:\CRF4\07312006\J560250.raw

```

263 tggagttccg cgttacataa cttacggtaa atggcccgcg tggctgaccg cccaacgacc 360
264 cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
265 attgacgtca atgggtggac tatttacggg aaactgcccc cttggcagta catcaagtgt 480
266 atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
267 atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
268 tcgctattac catgggtgatg cggtttttggc agtacatcaa tgggcgtgga tagcggtttg 660
269 actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
270 aaaatcaacg ggactttcca aaatgtcgtg acaactccgc ccattgacg caaatgggcg 780
271 gtaggcgtgt acgggtggag gtctatataa gcagagctct ctggctaact agagaaccca 840
272 ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gctggctagc 900
273 gtttaaactt aagcttggtg ccgagctcgg atccactctc ttccgcacgc ctgtctgcga 960
274 gggccagctg ttgggggtgag tactccctct gaaaagcggg catgacttct gcgctaagat 1020
275 tgtcagtttc caaaaacgag gaggatttga tattcacctg gcccgcggtg atgcctttga 1080
276 ggggtggccgc atccatctgg tcagaaaaga caatcttttt gttgtcaagc ttggtggcaa 1140
277 acgaccctga gagggcggtg gacagcaact tggcgatgga gcgcagggtt tgggttttgt 1200
278 cgcgatccgc gcgctccttg gccgcgatgt ttagctgcac gtattcgcgc gcaacgcacc 1260
279 gccattcggg aaagacgggt gtgcgctcgt cgggcaccag gtgcacgcgc caaccgcggt 1320
280 tgtgcagggt gacaagggtca acgctgggtg ctacctctcc gcgtaggcgc tcgttgggtcc 1380
281 agcagaggcg gccgcccttg cgcgagcaga atggcggtag ggggtctagc tgcgtctcgt 1440
282 cgggggggtc tgcgtccacg gtaaagaccc cgggcagcag gcgcgcgtcg aagtagtcta 1500
283 tcttgcatcc ttgcaagtct agcgctgct gccatgcgcg ggcggcaagc gcgcgctcgt 1560
284 atgggttgag tgggggaccc catggcatgg ggtgggtgag cgcggaggcg tacatgccgc 1620
285 aaatgtcgtg aacgtagagg ggctctctga gtattccaag atatgtaggg tagcatcttc 1680
286 caccgcggtg gctggcgcgc acgtaatcgt atagtctgtg cgagggagcg aggaggtcgg 1740
287 gaccgaggtt gctacgggcg ggctgctctg ctcggaagac tatctgcctg aagatggcat 1800
288 gtgagttgga tgatatgggt ggacgctgga agacgttgaa gctggcgtct gtgagacctg 1860
289 ccgcgtcagc cacgaaggag gcgtaggagt cgcgcagctt gttgaccagc tcggcggtga 1920
290 cctgcacgtc tagggcgcag tagtccaggg tttccttgat gatgtcatac ttatcctgtc 1980
291 cctttttttt ccacagctcg cggttgagga caaactcttc gcggtctttc cagtactctt 2040
292 ggatcggaac cccgtcggcc tccgaacgag atccgtactc cgcgcgcgag ggacctgagc 2100
293 gagtccgcat cgaccggatc ggaaaacctc tcgagaaagg cgtctaacca gtcacagtcg 2160
294 caagatccaa gatgaagcgc gcaagaccgt ctgaagatac cttcaacccc gtgtatccat 2220
295 atgacacgga aaccggctct ccaactgtgc cttttcttac tcctcccttt gtatccccc 2280
296 atgggtttca agagagtccc cctgggggtg tctcttttgc cctatccgaa cctctagtta 2340
297 cctccaatgg catgcttgcg ctcaaatgg gcaacggcct ctctctggac gaggccggca 2400
298 accttacctc caaaaatgta accactgtga gccacactct caaaaaaacc aagtcaaaca 2460
299 taaacctgga aatatctgca cccctcacag ttacctcaga agccctaact gtggctgccg 2520
300 ccgcacctct aatggtcgcg ggcaacacac tcacctgca atcacaggcc ccgctaaccg 2580
301 tgcacgactc caaacttagc attgccacce aaggacccct cacagtgtca gaaggaaagc 2640
302 tagccctgca aacatcagge cccctcacca ccacgatag cagtaccctt actatcactg 2700
303 cctcaccccc tctaactact gccactggta gcttgggcat tgacttgaaa gagccattt 2760
304 atacacaaaa tggaaaacta ggactaaagt acggggctcc tttgcatgta acagacgacc 2820
305 taaacacttt gaccgtagca actgggtccg gtgtgactat taataatact tccttgcaaa 2880
306 ctaaagttag tggagccttg ggttttgatt cacaaggcaa tatgcaactt aatgtagcag 2940
307 gaggactaag gattgattct caaaacagac gccttatact tgatgttagt tatccgtttg 3000
308 atgctcaaaa ccaactaaat ctaagactag gacagggcc tctttttata aactcagccc 3060
309 acaacttgga tattaactac aacaaaggcc tttacttggt tacagcttca aacaattcca 3120
310 aaaagcttga ggttaacctg agcactgcca aggggttgat gtttgacgct acagccatag 3180
311 ccattaatgc aggagatggg cttgaatttg gttcacctaa tgcaccaaac acaaatcccc 3240

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/560,250

DATE: 07/31/2006  
TIME: 08:48:17

Input Set : N:\RJAVED\10560250.txt  
Output Set: N:\CRF4\07312006\J560250.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:44; Xaa Pos. 4  
Seq#:45; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15  
Seq#:49; Xaa Pos. 4,7,9  
Seq#:50; N Pos. 1130,1157  
Seq#:52; N Pos. 1867,1875  
Seq#:56; N Pos. 1125

**VERIFICATION SUMMARY**

DATE: 07/31/2006

PATENT APPLICATION: US/10/560,250

TIME: 08:48:17

Input Set : N:\RJAVED\10560250.txt

Output Set: N:\CRF4\07312006\J560250.raw

L:15 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:838 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:842 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:31  
L:846 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:31  
L:1848 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:1852 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:44  
L:1853 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:44 after pos.:0  
L:1865 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:1869 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:45  
L:1873 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:45  
L:1877 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:45  
L:1878 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:45 after pos.:0  
L:1929 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:1933 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:49  
L:1937 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:49  
L:1938 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49 after pos.:0  
L:1950 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:1953 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:50  
L:1957 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:50  
L:2050 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50 after pos.:1112  
L:2117 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:2120 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:52  
L:2124 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:52  
L:2274 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52 after pos.:1845  
L:2370 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:2373 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:54  
L:2616 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:2619 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:56  
L:2623 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:56  
L:2716 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56 after pos.:1107